# Rapid evolution of male-biased gene expression in *Drosophila*

**Colin D. Meiklejohn*†, John Parsch‡, José M. Ranz*, and Daniel L. Hartl***

*Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138; and ‡Department of Biology II, Section of Evolutionary Biology, University of Munich, Luisenstrasse 14, 80333 Munich, Germany

Edited by Eviatar Nevo, University of Haifa, Haifa, Israel, and approved June 17, 2003 (received for review February 5, 2003)

**A number of genes associated with sexual traits and reproduction evolve at the sequence level faster than the majority of genes coding for non-sex-related traits. Whole genome analyses allow this observation to be extended beyond the limited set of genes that have been studied thus far. We use cDNA microarrays to demonstrate that this pattern holds in *Drosophila* for the phenotype of gene expression as well, but in one sex only. Genes that are male-biased in their expression show more variation in relative expression levels between conspecific populations and two closely related species than do female-biased genes or genes with sexually monomorphic expression patterns. Additionally, elevated ratios of interspecific expression divergence to intraspecific expression variation among male-biased genes suggest that differences in rates of evolution may be due in part to natural selection. This finding has implications for our understanding of the importance of sexual dimorphism for speciation and rates of phenotypic evolution.**

microarray | intraspecific variation | interspecific variation | cDNA

Anisogamous reproduction is common in many animal and plant species and can produce a number of conflicts with important evolutionary consequences. For example, differential selection coefficients between the two sexes can lead to stable genetic polymorphisms or a decline in population mean fitness (1). It can also drive accelerated rates of phenotypic evolution, as many morphologies associated with sex and reproduction diverge more rapidly than other phenotypes (2). Molecular techniques that provide rapid and quantitative measures of genotypic and phenotypic variation have extended this pattern to include accelerated rates of evolution among proteins with sexual or reproductive functions (3, 4). Since then, most data supporting this observation have come from homologous nucleotide sequences of genes that are associated with sex or reproduction. In ciliates, green algae, diatoms, angiosperms, fungi, and at least four animal phyla, unusually high ratios of nonsynonymous to synonymous substitutions ($d_N/d_S$) between species have been documented in sex-related genes (reviewed in ref. 5). Some of these genes also show high levels of intraspecific differentiation (5). In *Drosophila*, much of this work has focused on genes that are expressed in testes or accessory glands (e.g., refs. 6 and 7), although a high $d_N/d_S$ has also been observed for genes expressed in females and components of the sex determination pathway (8).

Protein coding sequences provide a natural context for studying rates of evolution, as the effect of a given nucleotide substitution on the polypeptide is predictable, and comparison between neighboring synonymous and nonsynonymous sites controls for mutation rate. Because of the lack of an analogous context for regulatory sequences, the rates and patterns of evolution in regions of the genome controlling gene expression are less well understood. Thus, it is not known whether the rapid rates of evolution among genes associated with sex and reproduction holds for gene expression as well. Because a large proportion of important phenotypic evolution may be the result of changes in gene expression (9, 10), understanding rates and patterns of regulatory change within and between species is critical for a comprehensive picture of biological evolution. Given the pattern seen for amino acid sequences and morphologies, we would predict that genes associated with sex should be evolving faster at the level of gene regulation as well. Indeed, much of the divergence among proteins in the male reproductive tract of *Drosophila* may be attributable to large changes in protein levels, which is likely due in part to changes in gene expression (3). To test this prediction, we obtained gene expression data for ≈1/3 of the genome from adult males of eight strains of *Drosophila melanogaster*, and from adult males and females of one strain of *D. melanogaster* and one strain of *Drosophila simulans*. By analyzing intra- and interspecific expression differentiation within males and the sex-specificity of expression in both species, we show that gene expression in males evolves more rapidly than in females. Genes that are male-biased in their expression have on average more intra- and interspecific divergence in expression than genes with female-biased expression. Furthermore, comparison of intra- and interspecific differentiation suggests that at least some of the excess in divergence among male-biased genes (MBGs) is due to differential selective pressures acting on the expression of different sex-biased classes of genes.

## Materials and Methods

**Fly Strains and cDNA Preparation.** Eight strains of *D. melanogaster* (three laboratory strains: Canton S, Oregon R, and Hikone R; an isofemale strain derived from St. Louis; and four lines derived from Zimbabwe: Zim53, Zim30, Zim29, and Zim2) were raised on standard medium at 25°C. Adult males were collected up to 24 h after eclosing, separated from females, and allowed to age an additional 3–4 days. Total RNA was extracted by using TRIzol reagent (Invitrogen) followed by chloroform extraction and isopropanol precipitation. Poly(A) RNA was purified by using the Oligotex Direct mRNA kit (Qiagen, Valencia, CA) and confirmed to be of high quality with a 2100 Bioanalyzer (Agilent Technologics, Palo Alto, CA). Two micrograms of poly(A) RNA was used as a template for SuperScript II reverse transcriptase (Invitrogen) in the presence of amino-allyl dUTP (Sigma). Cyanine-3 or cyanine-5 fluorochromes (Amersham Pharmacia) were incorporated after reverse transcription. Purification of cDNA and hybridizations were done following a published protocol (11). Labeled cDNAs were competitively hybridized to arrays by using the comparison scheme illustrated in Fig. 3, which is published as supporting information on the PNAS web site, www.pnas.org, with the following number of replicates per strain: Canton S, 13; Oregon R, 5; Hikone R, 3; St. Louis, 3; Zim53, 11; Zim30, 5; Zim29, 3; Zim2, 3.

**cDNA Microarrays.** A total of 5,928 clones from the *Drosophila* Gene Collection version 1.0 (12) were amplified by PCR with universal primers, and the products were confirmed by gel

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: MBGs, male-biased genes; FBGs, female-biased genes; UBGs, unbiased genes; OBGs, ovary-biased genes.

‡To whom correspondence should be addressed. E-mail: cmeiklejohn@oeb.harvard.edu.



www.pnas.org/cgi/doi/10.1073/pnas.1630690100

**Table 1. Overrepresentation of MBGs among genes with polymorphic expression within *D. melanogaster***

| Significance level of polymorphism | MBGs | FBGs | UBGs | *G* (2 df) |
|---|---|---|---|---|
| All genes | 910 | 1,499 | 1,397 | |
| $P < 0.05$ | 840 (92%) | 1,058 (71%) | 1,030 (74%) | 192 ($P < 0.001$) |
| $P < 0.01$ | 655 (72%) | 513 (34%) | 575 (41%) | 351 ($P < 0.001$) |

Subsets of genes include those that exhibit at least one pairwise difference between any two strains at the significance level indicated. *G*, *G* test of independence.

electrophoresis. Added to these was a set of 177 separately amplified controls, each of which was replicated from 1 to 16 times on the array. The PCR products were purified and mechanically spotted onto polylysine-coated glass slides (11). The results from these hybridizations have been deposited to the Gene Expression Omnibus (40) under accession nos. GPL356 and GSM7863–GSM7885.

**Statistical Analysis.** Relative gene expression levels were determined with a Bayesian method (Bayesian analysis of gene expression levels, BAGEL) (13) from the normalized ratio data (*Supporting Materials and Methods*, which is published as supporting information on the PNAS web site). This method estimates a normalized relative expression level for each strain and a single variance parameter across all strains from the Cy5/Cy3 ratios, on a gene-by-gene basis. It also calculates credible intervals from the stationary distribution of the Markov chain used to obtain the posterior distribution of the parameters (eight mean expression levels and one variance). For all pairwise intraspecific comparisons, unless otherwise stated, the threshold chosen for statistical significance was $P < 0.01$. This threshold signifies that the relative expression value for a given strain was greater than (or less than) that of another strain in >99% of the samples taken from the posterior distribution.

Sex-biased expression was defined by analysis of a parallel set of experiments comparing gene expression in adult males and females from a lab strain of *D. simulans* and the *D. melanogaster* strain Canton S (14). The significance threshold for sex-biased expression was defined by nonoverlapping 95% credible intervals, which was determined to be equivalent to $P < 0.00025$ by using a randomization approach (*Supporting Materials and Methods*). We define MBGs and female-biased genes (FBGs) as those with significantly different expression between the sexes (in the same direction) in both *D. melanogaster* and *D. simulans*. Unbiased genes (UBGs) are defined as those clones that show no significant expression between the two sexes in either species.

For a given gene, we describe intraspecific expression polymorphism ($S_e$) by the coefficient of variation of the relative expression levels among all eight strains. Similarly, the kurtosis for a given gene's expression ($K_e$) was calculated from the eight *D. melanogaster* expression values. Interspecific expression divergence ($D_e$) is described by the coefficient of variation between the mean of all eight *D. melanogaster* expression levels and the single *D. simulans* expression level (14) within a given sex. Differences in the distributions of these statistics are reported as the arithmetic mean of the statistic across all genes within a given sex-bias class (e.g., $\bar{S}_{eM}$, $\bar{S}_{eF}$, and $\bar{S}_{eU}$ for the mean expression polymorphism of MBGs, FBGs, and UBGs, respectively).

Relative expression levels and the associated credible intervals for each gene were recoded into discrete expression states by assigning to different states all strains for which the 95% credible intervals were nonoverlapping. Strains whose 95% credible intervals overlapped with multiple strains in different states were assigned to all of those states. The number of different transcriptional states found among the eight strains for a given gene was then tabulated.

Details regarding fluorescence ratio acquisition from microarray hybridizations, signal normalization and data quality control, assessment of false positive rates, and statistical analysis of previously published microarray data (15) are given in *Supporting Materials and Methods*.

## Results

The comparison scheme used here to obtain transcription profiles from adult males of eight strains of *D. melanogaster* is shown in Fig. 3. The strains were chosen to represent the range of genotypic and phenotypic variation known to exist in this species. Populations from Africa show significant differentiation from non-African populations in nucleotide variation (17) and mating behavior (18). Of the 4,905 clones selected for analysis, 2,289 showed differences that were significant between at least one pair of strains, whereas 297 are expected by chance (see *Supporting Materials and Methods*), indicating that at least 40% of the genome is detectably differentially regulated between males from interfertile populations of *D. melanogaster*. Pairs of strains showed from 218 to 928 genes with significantly different expression, where on average only 26 are expected by chance (Table 6, which is published as supporting information on the PNAS web site). This degree of differentiation in expression profile between strains is much higher than has been previously reported for *Drosophila* (16), and may reflect differing experimental designs and statistical methods as well as the inclusion in this study of the Zimbabwe strains. However, this level of variation is similar to the proportion of differentially expressed genes detected between two strains of *Saccharomyces cerevisiae* (19). Although the Zimbabwe strains do show evidence of differentiation from the Cosmopolitan strains in global gene expression (unpublished results), there are a surprisingly small number of genes that show fixed differences between the two groups (Table 7, which is published as supporting information on the PNAS web site).

The intraspecific comparisons were combined with information on interspecific divergence in gene expression between *D. melanogaster* and its sibling species, *D. simulans* (14). The combined data set consisted of 4,759 clones common to both experiments. As expected, the degree of differentiation among strains is smaller than the range of variation in transcription profiles seen between species. The variance in the distribution of $\log_2$ ratios across all genes between males of *D. melanogaster* and *D. simulans* is 0.208. Within *D. melanogaster*, the most divergent pair of strains has a variance in $\log_2$ ratios of 0.159, which is significantly lower than the interspecific comparison ($F_{4758,4758} = 1.30$, $P < 0.001$). Based on the distributions of $\log_2$ ratios, intraspecific differentiation ranges from 23% to 77% of interspecific divergence across the elements on these arrays.

There is a strong effect of sex-biased expression on intraspecific variation in gene expression, and this effect is reversed between MBGs and FBGs. Among genes that show significantly different expression between at least one pair of strains, there is a significant overrepresentation of MBGs and an underrepresentation of FBGs (Table 1). The strength of this effect increases as the stringency of the threshold chosen for statistical signifi-

EVOLUTION

**Table 2. Amounts and distribution of expression polymorphism is influenced by sex-biased expression**

| | MBGs | UBGs | FBGs | $P_{(M\ vs\ U)}$ | $P_{(U\ vs\ F)}$ |
|---|---|---|---|---|---|
| $\bar{S}_e$ | 0.158 | 0.140 | 0.108 | <0.001 | <0.001 |
| $\bar{K}_e$ | −0.086 | 0.327 | 0.229 | <0.001 | 0.121 |

*P* values were calculated from Wilcoxon rank-sum test.

**Table 3. Interspecific divergence is accelerated among MBGs when expressed in males**

| | $\bar{D}_{eM}$ | $\bar{D}_{eU}$ | $\bar{D}_{eF}$ | $P_{(M\ vs\ U)}$ | $P_{(U\ vs\ F)}$ |
|---|---|---|---|---|---|
| Expression in males | 0.164 | 0.129 | 0.120 | <0.001 | 0.055 |
| Expression in females | 0.132 | 0.125 | 0.121 | 0.150 | 0.727 |
| $P_{(expressing\ sex)}$ | <0.001 | 0.035 | 0.979 | | |

*P* values were calculated from a Wilcoxon rank-sum test.

cance is increased. By comparison with genes whose expression is not sex biased, the effect of sex bias on intraspecific transcriptional variation can be shown to be the result of both a reduction in variation among FBGs and an increase among MBGs; $\bar{S}_{eM}$ and $\bar{S}_{eF}$ are both significantly different from $\bar{S}_{eU}$ (Table 2).

Not only do MBGs on average have higher levels of expression polymorphism than FBGs or UBGs, but this variation is distributed differently among the eight strains than it is for the other two classes. The mean kurtosis ($\bar{K}_e$), for the expression levels of MBGs among the eight strains is significantly different from $\bar{K}_e$ for FBGs or UBGs; $\bar{K}_{eM}$ is more platykurtic, and the two other classes tend toward leptokurtosis (Table 2). A similar result is observed by assigning expression levels to discrete states, analogous to transcriptional alleles. The distributions of number of states for sex-biased and unbiased genes are shown in Fig. 1. MBGs have a much greater proportion of genes with two or more expression alleles than do the FBG or UBG classes, in which the majority of genes have a single expression state (male-biased vs. female-biased: $G = 388$, df = 3, $P < 0.001$). The distribution of expression states among FBGs is also significantly different from that of UBGs ($G = 52.8$, df = 3, $P < 0.001$). This result is not due to differences in ability to discriminate between states among the sex-bias classes. The average 95% credible interval across all eight strains is very similar for MBGs and FBGs (0.432 and 0.452, respectively), but is significantly higher in the UBGs (0.536; female-biased vs. unbiased: $t_s = -7.68$, df = 2481, $P < 0.001$). If this were responsible for the differences in transcription state distributions, it would result in an excess of monomorphic genes in the UBG class relative to the FBG class, but in fact the opposite result is observed (Fig. 1).

The interspecific hybridizations (14) allow a parallel set of observations to be made regarding the influence of sex-biased expression on interspecific divergence in gene expression. The results are consistent with the intraspecific data; $\bar{D}_e$ is significantly greater among MBGs than UBGs or FBGs (Table 3).

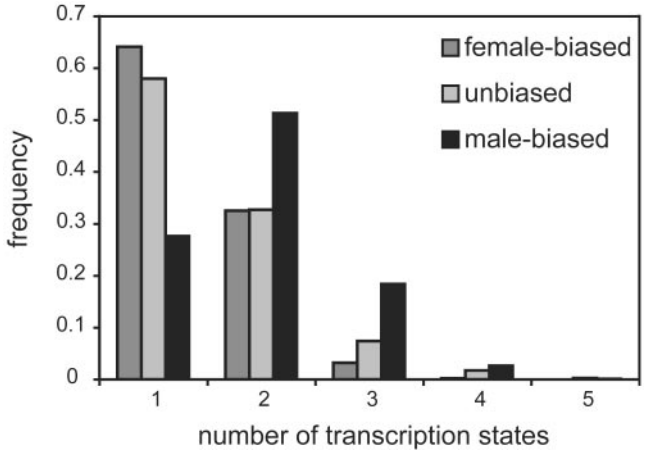


**Fig. 1.** Frequency distributions of gene expression states for male-enriched, female-enriched, and non-sex-biased genes. Relative gene expression levels were coded into discrete expression states as described in *Materials and Methods*.

These experiments also provide information on expression divergence for gene expression in females, as well as in males. Although the same pattern of increased divergence among MBGs relative to UBGs and FBGs exists for these genes when they are expressed in females, it is not nearly as strong, nor is it statistically significant (Table 3). Interestingly, MBGs are significantly more divergent when they are expressed in males than when they are expressed in females. Although not significant, this pattern holds for UBGs, suggesting that gene expression in males in general, and not just expression of MBGs, may be rapidly evolving.

Given DNA sequence data, deviations from the neutral expectation in the ratio of divergence to polymorphism have been used to infer the past activity of natural selection at a locus (20, 21). Although gene expression changes do not have a simple relationship with nucleotide sequence changes, a positive relationship between intra- and interspecific variation is predicted for neutrally evolving polygenic characters (22, 23), and has been empirically demonstrated for morphological traits that are presumably under selection as well (24). Elevated ratios of inter- to intraspecific variation in phenotypes associated with male reproduction have been used to infer the importance of directional selection on these characters relative to other types of morphologies (25). Thus, although we may not know *a priori* the neutral ratio of divergence to polymorphism for a given gene's expression, differences in this ratio between groups of genes may indicate disparate selective pressures acting on these groups. Fig. 2 shows $D_e$ plotted against $S_e$ for the three classes of sex bias. All three classes show a weak positive correlation between $S_e$ and $D_e$, and the correlation coefficient is significantly different between all three classes, indicating a different relationship of covariation between $S_e$ and $D_e$ for the three classes of sex bias.

The majority of sex-biased regulation in *Drosophila* has previously been shown to be the result of expression in germ-line tissues (15, 26); thus, transcription in testes and ovaries is most likely responsible for much of the sex-biased expression observed here, and in large part, it is genes expressed in the testes that are evolving rapidly and genes expressed in the ovaries that are evolving slowly. To confirm this conjecture, published data directly comparing *D. melanogaster* expression profiles of males and females, dissected testes and ovaries, and gonadectomized males and females (15) were analyzed and integrated with the results presented here. In addition to providing an independent identification of MBGs and FBGs, these data allow the description of analogous classes of genes defined by significantly different expression between testes and ovaries and between the somatic tissues of males and females.

The patterns of rapid expression evolution seen among MBGs relative to FBGs described above are also seen in the whole fly experiments of Parisi *et al.* (15). MBGs are overrepresented relative to FBGs among genes with polymorphic expression when the experiments of Parisi *et al*. (15) are used to determine sex-biased expression (data not shown). This same pattern is also found among genes with testis or ovary-biased expression (Table 4). Interestingly, this discrepancy is not observed for sex-biased genes when assayed in gonadectomized adults, because somatically MBGs and somatically FBGs show virtually identical

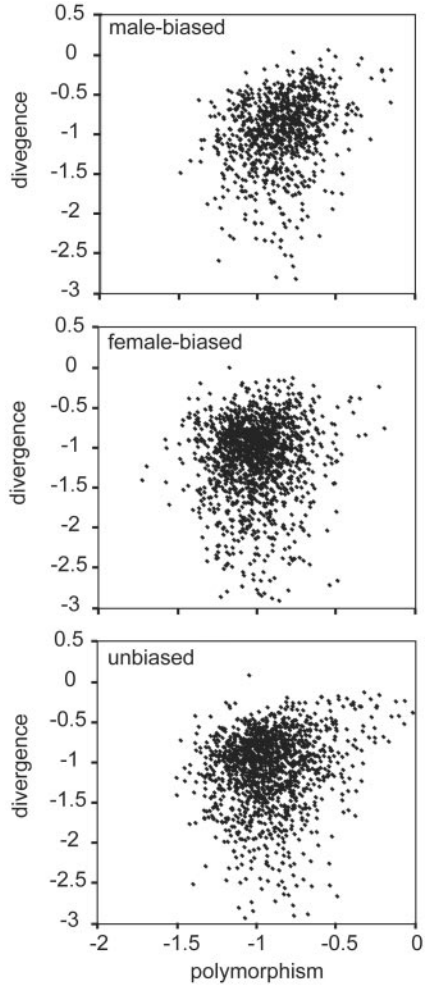


**Fig. 2.** $S_e$ and $D_e$ for male-biased, female-biased, and unbiased genes. The product moment correlation coefficients for each class are all significantly different from zero (MBG, $r = 0.273$, $P < 0.001$; FBG, $r = 0.064$, $P = 0.013$; UBG, $r = 0.162$, $P < 0.001$) and from each other (MBG vs. UBG, $z = 2.717$, $P = 0.007$; UBG vs. FBG, $z = 2.674$, $P = 0.008$). Both $S_e$ and $D_e$ were log-transformed before graphing.

representation among genes with polymorphic expression (Table 4). Furthermore, both of these classes appear to be overrepresented relative to genes with no somatic sex-biased expression, suggesting that sex-biased expression may be evolving relatively rapidly in somatic tissues as well as the gonads, but in both sexes.

**Table 4. Overrepresentation of testis-biased genes and somatically sex-biased genes among genes with polymorphic expression within *D. melanogaster***

| Significance level of polymorphism | TBGs | OBGs | sMBGs | sFBGs |
|---|---|---|---|---|
| All genes | 579 | 787 | 55 | 73 |
| $P < 0.05$ (%) | 543 (94)* | 579 (74)* | 49 (89) | 61 (84) |
| $P < 0.01$ (%) | 429 (74)* | 321 (41)* | 38 (69) | 49 (67) |

Significant departures from independence were determined by a *G* test with 1 df. Both sMBGs and sFBGs are overrepresented among polymorphic genes relative to genes with no somatic sex bias, of which 80% are polymorphic at $P < 0.05$ and 47% are polymorphic at $P < 0.01$. These results do not change if a different cutoff is chosen to assign significance to gonad or somatic sex-biased expression. TBGs, testis-biased genes; sMBGs, somatically MBGs; sFBGs, somatically FBGs. *, $P < 0.001$.

The summary statistics $\bar{S}_e$, $\bar{K}_e$, and $\bar{D}_e$ show similar patterns for gene expression variation in the gonads and soma as was described above for whole fly extractions (Table 5). Although greater intra- and interspecific variation is seen among genes with no sex-biased expression in the gonads compared with ovary-based genes (OBGs), this difference is not as significant as the difference between UBGs and FBGs seen from whole body extractions (compare Tables 2, 3, and 5). Among genes with somatic sex-biased expression, there is a greater $\bar{S}_e$ and $\bar{D}_e$ among both male-biased and female-biased genes than among genes with no somatic sex bias (Table 5). These patterns are also observed in the distributions of transcription states for testis-based genes, OBGs, somatically MBGs, and somatically FBGs (not shown).

A large proportion of the transcriptional differences observed between *D. melanogaster* and *D. simulans* involves the loss, gain, or reversal of sex-biased expression (14). Examination of intra- and interspecific expression variation in these genes does not reveal as clear a pattern as that observed for genes retaining an ancestral sex bias in both *D. melanogaster* and *D. simulans.* This is most likely due to a diversity of selective forces acting on genes with rapidly evolving sex bias. However, such genes do appear to be more variable in their expression within *D. melanogaster* than those that have retained the ancestral sex bias, as shown by a greater $\bar{S}_e$ among genes that are sex biased in one species only than among those that are sex biased in both species (male-biased in one species only, $\bar{S}_e = 0.180$, $P > 0.05$, Wilcoxon rank-sum test; female-biased in one species only, $\bar{S}_e = 0.122$, $P < 0.001$, Wilcoxon rank-sum test). However, among genes with a novel sex bias, there is still a correlation between $S_e$ and sex bias, as genes that are male-biased in *D. melanogaster* or *D. simulans* only have a $\bar{S}_e$ that is significantly greater than genes that are female-biased in one species only ($P < 0.001$, Wilcoxon rank-sum test).

## Discussion

The data presented here indicate that rates of both intraspecific and interspecific differentiation of gene expression in *Drosophila* are correlated with sex-biased expression, and that this difference is largely a function of gene expression in testes and ovaries. Furthermore, among somatically expressed genes, sex-biased expression in both sexes appears to evolve more rapidly than sexually monomorphic expression. Analogous results have come from morphological studies in *Drosophila* that documented a higher rate of intra- and interspecific divergence among morphologies associated with male reproduction than nonreproductive morphologies (25). These conclusions are not the result of nucleotide sequence divergence within or between species causing spurious inferences of changes in gene expression. Data from competitive hybridizations using genomic DNA extracted from *D. melanogaster* and *D. simulans* indicate that sequence divergence between these two species has a small effect on hybridization signal intensity that is within the range of experimental error associated with cDNA hybridizations (14). Furthermore, across the clones on these arrays, a greater number of MBGs are found in *D. simulans* than in *D. melanogaster* (14), which cannot be the result of sequence divergence. Consistent with our results on rates of expression evolution, a subset of male germ-line genes in *Drosophila* are known to be enriched for sequences with no detectable homologs in other eukaryotic genomes (26), suggesting that MBGs may be on the whole younger than other classes of genes. The accelerated rate of evolution among MBGs may therefore extend further back in time than the comparison between *D. melanogaster* and *D. simulans*, and is consistent with the hypothesis that evolution of male-specific phenotypes may be often driven by the creation of new genes (e.g., refs. 27 and 28).

One interpretation of these results is that mutations affecting the expression of MBGs on average experience greater (i.e.,

**Table 5. Expression polymorphism and divergence as a function of sex-biased expression in the gonads and soma**

| | TBGs | UBGs | OBGs | $P_{(T\ vs\ U)}$ | $P_{(U\ vs\ O)}$ | sMBGs | sUBGs | sFBGs | $P_{(M\ vs\ U)}$ | $P_{(M\ vs\ F)}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $\bar{S}_e$ | 0.162 | 0.124 | 0.119 | <0.001 | 0.175 | 0.202 | 0.131 | 0.221 | <0.001 | 0.280 |
| $\bar{K}_e$ | −0.229 | 0.248 | 0.191 | <0.001 | 0.486 | −0.090 | 0.182 | 0.174 | 0.193 | 0.115 |
| $\bar{D}_e$ | 0.180 | 0.134 | 0.125 | <0.001 | 0.225 | 0.207 | 0.135 | 0.224 | <0.001 | 0.536 |

*P* values were calculated from a Wilcoxon rank-sum test.

either more positive or less negative) selection coefficients than UBGs or FBGs. A larger positive average selection coefficient would result in the fixation of a greater number of beneficial mutations affecting gene expression of MBGs, whereas less negative selection coefficients would result in a smaller fraction of deleterious mutations contributing to intraspecific variation affecting the expression of MBGs. Both of these hypotheses are suggested in the greater correlation between $S_e$ and $D_e$ seen among MBGs (Fig. 2). Differences in this relationship between the sex-biased classes are largely due to an elevated $D_e/S_e$ ratio among the most extremely male-biased genes. The 324 genes with the most significantly testis-biased expression ($P < 0.001$) have a higher average $D_e/S_e$ than the 290 corresponding OBGs (Wilcoxon rank sum test, $P = 0.04$). Circumstantial evidence for the role of positive selection in the differentiation among MBGs is seen in the patterns of divergence between *D. melanogaster* and *D. simulans* (Table 3). If relaxed selection is driving the divergence of MBGs, the fact that gene expression among MBGs is less divergent in females than it is in males requires that the neutral mutations that fix and cause changes in expression have their effects in males only, because their regulation is more conserved in females.

An alternate interpretation of these data are that there is a fundamentally different relationship between fold-change in expression and effect on fitness between the sex-biased classes of genes. Because gene regulation in male gametogenesis appears to be highly specialized in both mammals and insects (29), we might expect the evolution of gene expression in testes to be unusual. However, this explanation does not address the differences observed here between FBGs and UBGs (i.e., Tables 1 and 2).

It is important to remember that the intraspecific comparisons presented here include gene expression data from males only. However, preliminary results indicate that the relationship between sex bias and rates of intraspecific expression evolution seen here for gene expression in males holds for expression in females as well. Assessment of expression profiles in virgin females of Canton S and Zim2 using the same methods described above reveals a significant overrepresentation of MBGs among those genes with significant differences between these two strains. Furthermore, in these females $\bar{S}_{eM} > \bar{S}_{eU} > \bar{S}_{eF}$, and these differences are highly significant (J.M.R., unpublished data). This finding is in agreement with the results from the interspecific comparisons, which show that FBGs show similar or reduced amounts of interspecific differentiation than MBGs when these genes are expressed in females (Table 3).

The literature documenting elevated rates of evolution among genes associated with reproduction has included numerous examples of genes with functions in both males and females (4, 5). This is in contrast to the results presented here, which show that the expression of FBGs evolves more slowly than that of both MBGs and UBGs, although this pattern is not as extreme as the accelerated evolution observed for male-biased expression. Some of the studies that have found rapidly evolving genes associated with female reproduction have focused on a small sample for which there was an *a priori* expectation of positive selection (e.g., ref. 30, but see ref. 4). The low correlation between $S_e$ and $D_e$ for all three sex-bias classes shows that the rate of expression evolution for any given gene is likely to be idiosyncratic, and we observe a number of FBGs with high levels of intra- and interspecific expression variation. Nonetheless, assaying ≈1/3 of the *Drosophila* genome indicates that rapid evolution of expression is far more prevalent among MBGs than FBGs, and indicates the value of data sets of this size.

One potential source of error in the above analyses is the assumption of independence across genes in their intra- and interspecific variation in expression. This assumption will be violated when a single genetic locus influences variation in the expression of multiple genes simultaneously, as will result from coordinate regulation. In the most extreme scenario, all genes expressed in the testes might appear to be up-regulated in a given strain of *D. melanogaster* because of an increase in the relative size of the testes in that strain. Such an extreme bias can be ruled out by the large number of genes that show all possible patterns of covariation among these strains, indicating that at a broad scale there are many groups of independently regulated genes. Further argument to this point can be made by reference to the few experiments to date that have examined the genetics of gene expression. Studies in *S. cerevisiae* (19), mice, and maize (31) indicate that 35–80% of QTLs that influence expression of a gene map to the gene itself, suggesting *cis*-regulation. The fraction of *cis*-acting genetic factors increases with more stringent statistical cutoffs (31), suggesting that large changes in expression may more often be in *cis*, whereas *trans*-acting mutations are more often of small effect. Although these numbers are influenced by the power of each experimental design, these data from three different biological kingdoms suggest that a large fraction (>30%) of large effect variants affecting gene expression are in *cis*, and that this may be a phenomenon intrinsic to eukaryotic gene expression. This does not address how the remaining fraction of variation in gene expression is distributed across many unlinked factors of small effect. Nonetheless, it is unlikely that these caveats could affect the nature of our conclusions, or render them statistically insignificant. For example, all of the comparisons made in Tables 1 and 4 remain significant at $P < 0.01$, when the numbers of genes are multiplied by a factor of 0.1 (as might be appropriate if, on average, a single genetic variant were responsible for changes in the expression of 10 downstream genes). However, this issue will require population genetic studies of the inheritance of global gene expression to be conclusively addressed.

It is tempting to speculate that MBG expression may be related to the evolution of hybrid male sterility (32). The genetic factors that influence male fertility appear to evolve much faster at both intra- and interspecific levels than those influencing female fertility or viability in either sex in *Drosophila*. This rapidity is evident in the disproportionately high amounts of genetic variation affecting male fertility observed in mutation–accumulation lines (33) and the excess of hybrid male sterility factors relative to hybrid female sterility factors that have accumulated between closely related species (34–

36). None of these patterns is caused by an excess of loci affecting male fertility, because mutagenesis screens indicate that approximately seven times more genes influence viability than male fertility, whereas similar numbers affect fertility in the two sexes (37). A causal relationship between these observations would require that the rapid evolution of gene expression in MBGs leads to the misexpression of these genes in sterile hybrid males. A recent study of gene expression in hybrids between *Drosophila mauritiana* and *D. simulans* found that MBGs were preferentially misexpressed in the sterile $F_1$ males (38), lending support to this hypothesis. Together, these patterns of gene expression and misexpression are consistent with the idea that rapid evolution of male reproductive characters contributes to Haldane's rule (39) for hybrid male sterility in *Drosophila*.

We thank the members of the Harvard *Drosophila* Microarray Consortium and the Bauer Center for Genomics Research for help in creating the arrays; the Hartl and Wakeley laboratories, A. M. Wilczek, and especially J. P. Townsend for helpful discussions; and Rama Singh, Mohamed Noor, and an anonymous reviewer for valuable comments on the manuscript. This work was supported by National Institutes of Health Grant GM60035 (to D.L.H.) and a postdoctoral fellowship from the Ministerio de Ciencia y Tecnología (to J.M.R.).

1. Ewens, W. J. (1979) *Mathematical Population Genetics* (Springer, Berlin).
2. Darwin, C. (1871) *The Descent of Man, and Selection in Relation to Sex* (John Murray, London).
3. Coulthart, M. B. & Singh, R. S. (1988) *Mol. Biol. Evol.* **5,** 182–191.
4. Civetta, A. & Singh, R. S. (1995) *J. Mol. Evol.* **41,** 1085–1095.
5. Swanson, W. J. & Vacquier, V. D. (2002) *Annu. Rev. Ecol. Syst.* **33,** 161–179.
6. Swanson, W. J., Clark, A. G., Waldrip-Dail, H. M., Wolfner, M. F. & Aquadro, C. F. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 7375–7379.
7. Parsch, J., Meiklejohn, C. D., Hauschteck-Jungen, E., Hunziker, P. & Hartl, D. L. (2001) *Mol. Biol. Evol.***18,** 801–811.
8. Civetta, A. & Singh, R. S. (1998) *Mol. Biol. Evol.* **15,** 901–909.
9. Britten, R. J. & Davidson, E. H. (1969) *Science* **165,** 349–357.
10. Carroll, S. B., Grenier, J. K. & Weatherbee, S. D. (2001) *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design* (Blackwell Science, Malden, MA).
11. Eisen, M. B. & Brown, P. O. (1999) *Methods Enzymol.* **303,** 179–205.
12. Rubin, G. M., Hong, L., Brokstein, P., Evans-Holm, M., Frise, E., Stapleton, M. & Harvey, D. A. (2000) *Science* **287,** 2222–2224.
13. Townsend, J. P. & Hartl, D. L. (2002) *Genome Biol.* **3,** RESEARCH0071.1–0071.16.
14. Ranz, J. M., Castillo-Davis, C. I., Meiklejohn, C. D. & Hartl, D. L. (2003) *Science* **300,** 1742–1745.
15. Parisi, M., Nuttall, R., Naiman, D., Bouffard, G. G., Malley, J., Andrews, J., Eastman, S. & Oliver, B. (2003) *Science* **299,** 697–700.
16. Jin, W., Riley, R. M., Wolfinger, R. D., White, K. P., Passador-Gurgel, G. & Gibson, G. (2001) *Nat. Genet.* **29,** 389–395.
17. Begun, D. J. & Aquadro, C. F. (1993) *Nature* **365,** 548–550.
18. Hollocher, H., Ting, C.-T., Pollack, F. & Wu, C.-I. (1997) *Evolution (Lawrence, Kans.)* **51,** 1175–1181.
19. Brem, R. B., Yvert, G., Clinton, R. & Kruglyak, L. (2002) *Science* **296,** 752–755.
20. Hudson, R. R., Kreitman, M. & Aguadé, M. (1987) *Genetics* **116,** 153–159.
21. McDonald, J. H. & Kreitman, M. (1991) *Nature* **351,** 652–654.
22. Lande, R. (1976) *Evolution (Lawrence, Kans.)* **30,** 314–334.
23. Lynch, M. & Hill, W. G. (1986) *Evolution (Lawrence, Kans.)* **40,** 915–935.
24. Kluge, A. G. & Kerfoote, W. C. (1973) *Am. Nat.* **107,** 426–442.
25. Civetta, A. & Singh, R. S. (1998) *Evolution (Lawrence, Kans.)* **52,** 1080–1092.
26. Arbeitman, M. N., Furlong, E. E. M., Imam, F., Johnson, E., Null, B. H., Baker, B. S., Krasnow, M. A., Scott, M. P., Davis, R. W. & White, K. P. (2002) *Science* **297,** 2270–2275.
27. Nurminsky, D. I., Nurminskaya, M. V., De Aguiar, D. & Hartl, D. L. (1998) *Nature* **396,** 572–575.
28. Long, M. & Langley, C. H. (1993) *Science* **260,** 91–95.
29. Kleene, K. C. (2001) *Mech. Dev.* **106,** 3–23.
30. Swanson, W. J., Yang, Z., Wolfner, M. F. & Aquadro, C. F. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 2509–2514.
31. Schadt, E. E., Monks, S. A., Drake, T. A., Lusis, A. J., Che, N., Colinayo, V., Ruff, T. G., Milligan, S. B., Lamb, J. R., Cavet, G., *et al.* (2003) *Nature* **422,** 297–302.
32. Laurie, C. C. (1997) *Genetics* **147,** 937–951.
33. Fry, J. D., Heinsohn, S. L. & Mackay, T. F. C. (1998) *Genetics* **148,** 1171–1188.
34. True, J. R., Weir, B. S. & Laurie, C. C. (1996) *Genetics* **142,** 819–837.
35. Wu, C.-I. & Davis, A. W. (1993) *Am. Nat.* **142,** 187–212.
36. Tao, Y. & Hartl, D. L. (2003) *Evolution (Lawerence, Kans.),* in press.
37. Lindsley, D. L. & Tokuyasu, K. T. (1980) in *The Genetics and Biology of Drosophila*, eds. Ashburner, M. & Wright, T. R. F. (Academic, New York), Vol. 2d, pp. 226–294.
38. Michalak, P. & Noor, M. A. F. (2003) *Mol. Biol. Evol.* **20,** 1070–1076.
39. Haldane, J. B. S. (1922) *J. Genet.* **12,** 101–109.
40. Edgar, R., Domrachev, M. & Lash, A. E. (2002) *Nucleic Acids Res.* **30,** 207–210.